JOURNAL OF BASIC SCIENCE AND ENGINEERING

A COMBINED DEEP REINFORCEMENT AND SUPERVISED LEARNING TECHNIQUE TO IDENTIFY CYBER-ATTACKS IN DISTANCE RELAYS

Srishty[1], Uday Patil[2]
[1]PG Scholar, Central University of Karnataka, Kalaburagi
[2]Assistant Professor, Central University of Karnataka, Kalaburagi

ABSTRACT

This paper proposes a Multi Agent Deep-reinforcement learning algorithm for detecting cyber-attacks in distance relays. Distance relays are widely used in the power grid for protection against faults, but they are vulnerable to cyber-attacks due to their distributed nature. The proposed algorithm uses multiple agents to learn an optimal policy for detecting cyber-attacks in distance relays. Each agent is trained using a combination of deep reinforcement learning and supervised learning techniques. The agents are trained to identify attacks by observing the current state of the system and taking actions that optimize a reward function. The reward function is designed to maximize the detection accuracy of the agents while minimizing the false alarm rate. The algorithm is evaluated on a benchmark dataset of simulated cyber-attacks. The results show that the proposed algorithm outperforms existing approaches in terms of attack detection accuracy and false alarm rate.

Keywords: ***Artificial Intelligence, Reinforcement Learning, Deep Q-Network, Cyber-attacks***

1. INTRODUCTION

Distance relays are designed to detect faults at a distance from the point at which the fault occurred, thus providing faster and better protection. Distance relays use the principle of impedance to detect faults. Impedance is the resistance an electric current encounter when traveling through a power system. When a fault occurs, the impedance between the fault point and the relay increases significantly, causing the relay to trip. Deep reinforcement learning and industrial control elements of networks are extensively employed for developing a unique reward and learning mechanism [1]-[15]. Deep reinforcement learning and industrial control elements of networks are extensively employed for developing a unique reward and learning mechanism. Deep reinforcement learning is also used to develop a technique for identifying abnormalities in industrial control systems.

Several methods for identifying cyber-attacks in remote relays, the system that has been proposed which employs a large number of agents [16]-[18]. Diverse training techniques, including as deep reinforcement learning, supervised learning, and unsupervised learning, are

used to instruct each agent. Throughout their training, the agents engage in activities that are intended to optimize the value of a reward function in order to identify attacks. They also keep an eye on the system's condition. The incentive function was created to decrease the amount of false positives encountered by the agents and increase their detection accuracy. Analyzing the recommended strategy involves using a standard dataset made up of simulated cyber-attacks. Due to the advent of high-speed communication services, power grids have evolved from traditional transfer networks to cyber-physical platforms, posing significant security challenges [18]-[21]. Confidentiality, integrity, and availability are three key concepts for the security of smart grids presented by the National Institute of Standards and Technology (NIST). Multi-Agent Distributed Deep Learning (MADDL) is the most advisable to counter cyberattacks in long- distance relays [22]-[24]. Using graph theory, transformation of the protective system with several distant relays into a decentralised multi-agent system can be achieved. Each agent is believed to have its deep neural network for detecting threats, and these agents are interconnected so that they may share voltage and current data. Train data is used to fine-tune the detection structures, which are then subjected to testing on a test dataset. The suggested technique has been shown to detect more than 99.88% of errors and cyberattacks in simulations. The distance relays in a power grid form a communication network, with each relay only linked to its immediate neighbours. High-quality performance in detecting cyber-attacks and easy compatibility with the expansion of power grids are achieved by MADDL algorithm despite its requirement for data for tweaking and a bit complicated implementation [23]-[24].

The goal of reinforcement learning (RL), a subfield of AI, is to programmatically learn the best possible decisions over time. It brings AI closer to social awareness by including time as a new component in the learning process. Several technological and industrial fields have taken up the RL challenge [25]. To determine which lines are most at risk of failing in the event of a coordinated multistage assault, this study offers an approach based on multi-agent deep reinforcement learning and prioritised experience replay. It creates a defense plan based on the ideal offensive line sequences and the defense capabilities. The suggested algorithm and the intended approach were shown to be successful in simulations, allowing the power system operator to more efficiently employ their limited defense resources and lessen the damage done by coordinated multistage assaults .

Deep reinforcement learning as a technique for actively mitigating the consequences of a cyberattack on a portion of the DER units in the network. This might be done by figuring out the best settings for the control logic of a group of uncompromised DER units. This shift is being accomplished via standardizing the functionality of grid standards or international norms. These two strategies are both in use. DER, on the other hand, are distinct in that distribution utilities do not normally own or run them; as a result, they pose a new emergent threat vector for cyber-physical assaults. Combining a deep feed forward neural network method and a reinforcement learning method based on Q-learning yields a new generation of network intrusion detection tools.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

This technique integrates a network intrusion detection technique with a reinforcement learning technique based on Q-learning [26]-[27]. His urged Deep Q-Learning (DQL) paradigm provides a network framework with continuing auto-learning capability. By utilizing an automated process of trial-and-error, this feature enables the network environment to recognize various types of network intrusions while simultaneously improving its detection capabilities. The results of his experiments demonstrate that his suggested DQL is superior to existing methods of machine learning that are analogous in nature and works very well when it comes to identifying various types of intrusions. Using deep learning and reinforcement learning, the authors of the paper have created an artificial intelligence-based strategy for process control. This framework encourages the creation of control strategies that may gain knowledge via direct interaction with a plant's output. In games and physical activities, human level control has been reached by fusing deep learning with reinforcement learning [28]-[31].

In this paper, the method for evaluating cybersecurity that is intended to evaluate the cyber physical security of EPS using Deep-Reinforcement learning (DRL) is chosen and focuses on deep learning approach to detect malicious assaults on SCADA systems, by using Deep reinforcement learning, which acquires information from sensor data and identifies patterns indicative of defects, can be used to detect and diagnose errors in industrial processes. The formed methodology has been used to power grids with 6, 14, and 118 buses in three case studies.

2. PROPOSED SYSTEM

The well-known IEEE 14 bus structure is a benchmark power system that is put to use for the purpose of putting several power system analysis and optimization methods for the test. In the 1970s, the Institute of Electrical and Electronics Engineers (IEEE) established it in order to investigate the effectiveness of several algorithms for power system analysis. These algorithms include load flow analysis, voltage stability analysis, and contingency analysis.

The IEEE 14 bus system has 11 loads, five generators, and 14 buses altogether. The buses are linked together by means of twenty transmission cables and six transformers.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

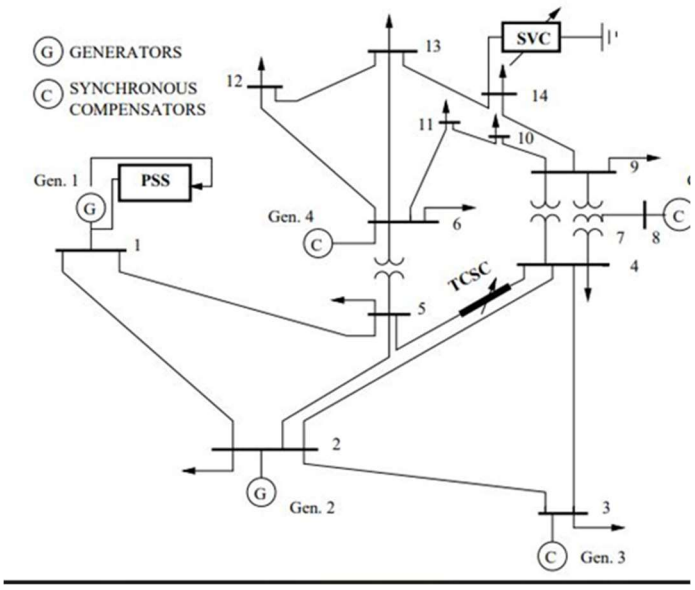

Fig.1: Electrical Line Diagram of IEEE 14 Bus System

The generators are modeled as voltage-controlled buses, and the amplitude of the voltage as well as the phase angle are held constant at each generating bus. The loads are represented as constant-power loads, which means that the amount of actual power and reactive power consumption at each load bus is held constant. In both academic and commercial settings, the IEEE 14 bus system has been put to extensive use for the purposes of testing and verifying various power system analysis and optimization techniques. It is a common system that is used as a benchmark for comparing the performance of various algorithms and methodologies, as well as for analyzing the influence that various parameters have on the dependability and stability of power systems.

To implement the proposed work, required following facilities.

MATLAB and Simulink: In order to use these programs, you will need to have MATLAB and Simulink already installed on your computer. MATLAB is a programming language that is used for mathematical calculation and data analysis.

Deep learning toolbox: This may be done through the use of MATLAB's deep learning toolbox. To work with deep reinforcement learning, you will need to have this toolbox already installed on your system.

Data: In order to train your deep reinforcement learning agent, you are going to need some data. This may use computer simulations, data collected from the actual world, or a mix of the two.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

3. TRAINING THE SYSTEM WITH DATASET

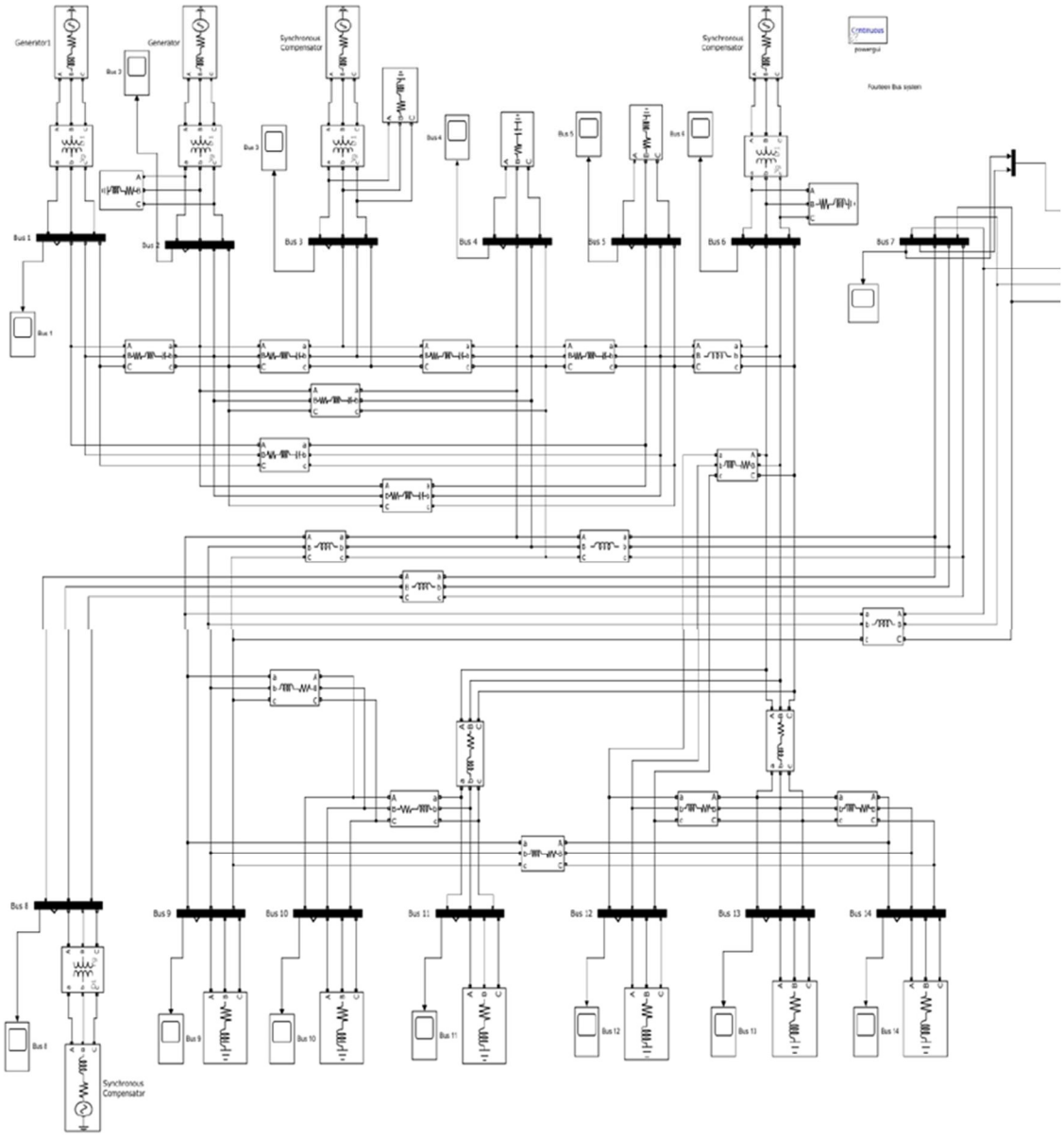

Fig.2: IEEE 14 BUS simulation model

Fig. 2 states that the simulation model of IEEE 14 Bus protocol which is used to extract data for deep reinforcement learning algorithm. The necessary data for the IEEE 14-bus system is collected and entered in the model. This includes parameters such as bus voltages, line impedances,

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

generator details, and load data. MATLAB variables were initialized with the collected data and stored in MATLAB arrays or structures for easy access. Analysis was carried out to determine the steady-state operating conditions of the system.

Time	V1	V2	V3	C1	C2	C3	Class	action-0	action-1
0	-0.15737	-0.76027	0.917636	-1.8E-09	-1.1E-09	2.9E-09	0	1	-1
7.94E-07	-0.15712	-0.76042	0.917549	-1.8E-09	-1.1E-09	2.9E-09	0	1	-1
2.38E-06	-0.15664	-0.76073	0.917376	-1.8E-09	-1.1E-09	2.9E-09	0	1	-1
1.03E-05	-0.15423	-0.76228	0.916503	-1.8E-09	-1.1E-09	2.9E-09	0	1	-1
4.19E-05	-0.14459	-0.76838	0.912971	-1.8E-09	-1.1E-09	2.91E-09	0	1	-1
0.000163	-0.1077	-0.79097	0.898666	-1.7E-09	-1.2E-09	2.92E-09	0	1	-1
0.000283	-0.07065	-0.81242	0.883069	-1.6E-09	-1.3E-09	2.93E-09	0	1	-1
0.000404	-0.0335	-0.83271	0.866202	-1.5E-09	-1.4E-09	2.93E-09	0	1	-1
0.000525	0.003702	-0.85179	0.848092	-1.4E-09	-1.5E-09	2.94E-09	0	1	-1
0.000616	0.03201	-0.8655	0.833491	-1.4E-09	-1.6E-09	2.93E-09	0	1	-1
0.000708	0.060291	-0.87849	0.818198	-1.3E-09	-1.6E-09	2.93E-09	0	1	-1
0.000773	0.080151	-0.88718	0.807034	-1.2E-09	-1.7E-09	2.92E-09	0	1	-1
0.000837	0.099978	-0.89552	0.795538	-1.2E-09	-1.7E-09	2.92E-09	0	1	-1
0.000923	0.126115	-0.90596	0.779844	-1.1E-09	-1.8E-09	2.91E-09	0	1	-1
0.001008	0.152162	-0.91575	0.763589	-1E-09	-1.9E-09	2.89E-09	0	1	-1
0.001093	0.178099	-0.92488	0.746786	-9.5E-10	-1.9E-09	2.88E-09	0	1	-1
0.001179	0.203909	-0.93335	0.729446	-8.8E-10	-2E-09	2.86E-09	0	1	-1
0.001247	0.224417	-0.93964	0.71522	-8.2E-10	-2E-09	2.85E-09	0	1	-1
0.001315	0.244822	-0.94549	0.700667	-7.6E-10	-2.1E-09	2.83E-09	0	1	-1
0.001383	0.265115	-0.95091	0.685792	-7E-10	-2.1E-09	2.82E-09	0	1	-1
0.001451	0.285287	-0.95589	0.670602	-6.3E-10	-2.2E-09	2.8E-09	0	1	-1
0.00152	0.305328	-0.96043	0.655106	-5.7E-10	-2.2E-09	2.78E-09	0	1	-1
0.001588	0.325228	-0.96454	0.639309	-5.1E-10	-2.2E-09	2.76E-09	0	1	-1
0.001656	0.34498	-0.9682	0.623219	-4.5E-10	-2.3E-09	2.74E-09	0	1	-1
0.001724	0.364573	-0.97142	0.606843	-3.9E-10	-2.3E-09	2.71E-09	0	1	-1
0.001812	0.389549	-0.9749	0.585353	-3.1E-10	-2.4E-09	2.68E-09	0	1	-1
0.0019	0.414228	-0.97765	0.563419	-2.3E-10	-2.4E-09	2.65E-09	0	1	-1
0.001987	0.438593	-0.97965	0.541056	-1.4E-10	-2.5E-09	2.61E-09	0	1	-1
0.002075	0.462625	-0.98091	0.518283	-6.4E-11	-2.5E-09	2.57E-09	0	1	-1
0.002163	0.486305	-0.98142	0.495116	1.72E-11	-2.6E-09	2.53E-09	0	1	-1

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

Table 1: Glimpse of dataset

The Table 1 state that how data is collected for training and testing of deep-learning algorithm model in this we have 6 states (V1, V2, V3, C1, C2, C3) and two action which is 0 and 1 ,such that 0 is for fault detection and 1 for cyber-attack detection.

MATLAB simulation of IEEE 14 bus was carried out and then afterwards deep reinforcement learning algorithm was applied in which we have set initial values of Q table with zeros or random values and then set learning rate (alpha=0.9), gamma=0.75 etc. Thereafter, to set number of Episodes and for each Episodes, we have to reset environment to initial state. For each step in the episode, choose an using epsilon-greedy rule and perform the chosen action in the environment afterwards observe new state and reward.

Now move towards updating the Q table using Q-learning Equation:

Q (current_state,next_state)= Q(current_state,next_state)+alpha *TD

Then transition to next state and if the episode is completed then End training process of deep reinforcement model.

Learned final Q- table to exploit the environment and choose an action with the highest value for the current state is used to train the data. After training 70% of data as shown, we have to involve in Testing and Evaluation process with 30% of the data

Final Q - table.	
8.0000	10.0000
0.1835	10.0000
-2.553	3.3025
-0.1000	0.4128
-0.1000	0
0	0

Table 2: Q-state dataset

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

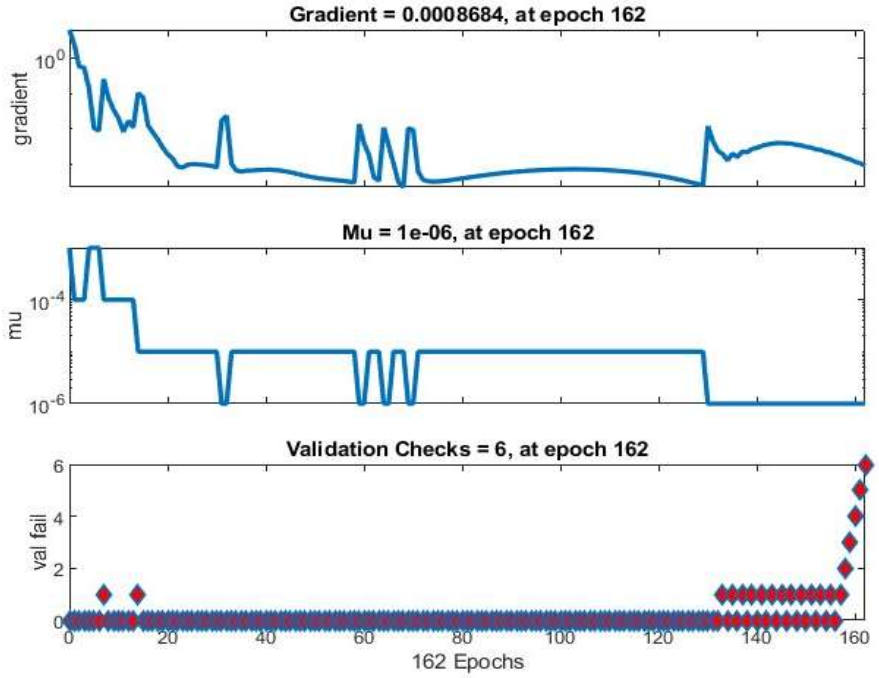

Fig. 3: Training state

4. RESULTS

From the below confusion matrix tables there are 2 classes, class-0 represents the Not a cyber-attack but fault and class-1 represents the Not a fault but cyber-attack.

4.1 Testing

A. Case Study for IEEE-14 Bus Train

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

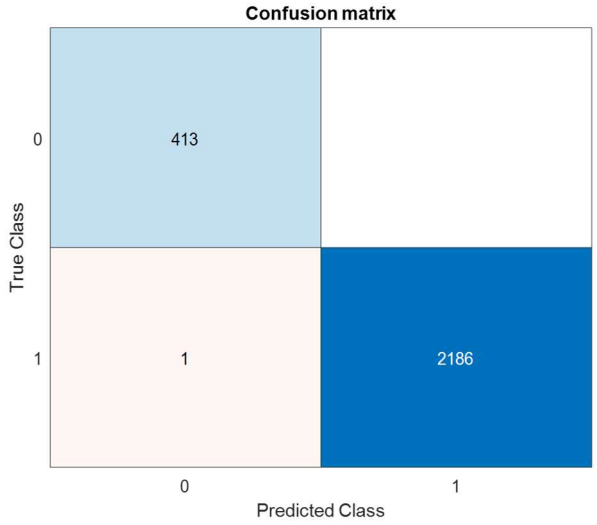

Table 2 Confusion Matrix Table (Train)

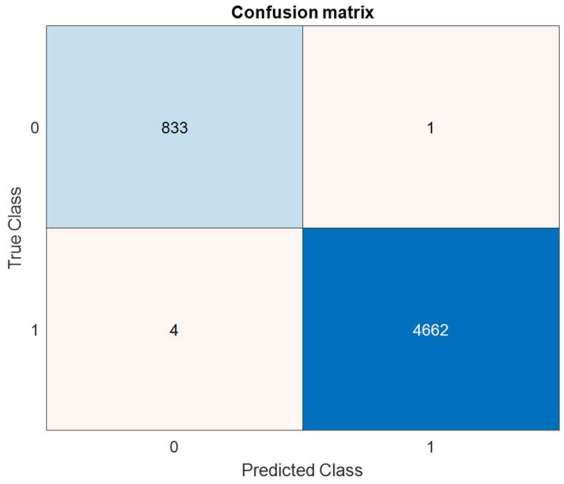

Table 3 Confusion Matrix Table (Test)

From the above table 4 of TRAIN-SET of IEEE-14 BUS, it is observed that there are 413 faults are correctly classified out of 413 faults and we can also detect that there are 2186 cyber-attacks are correctly identified out of 2187 cyber-attacks

From the above table 5 of TEST-SET of IEEE-14 BUS, it is observed that there are 833 faults are correctly classified out of 834 faults and we can detect that there are 4662 cyber-attacks are

correctly identified out of 4666 cyber-attacks.

B. Case Study-III for IEEE-118 Bus Train

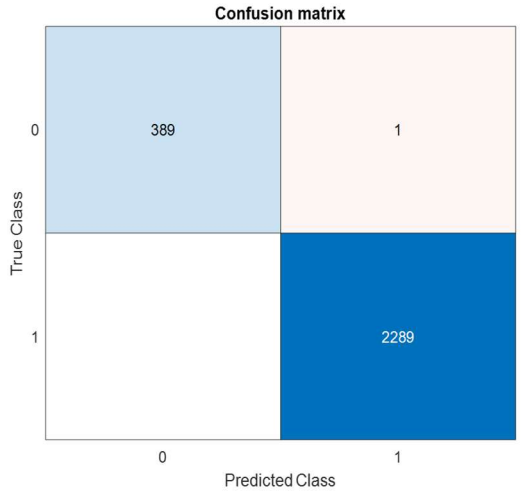

Table 4 Confusion Matrix Table (Train)

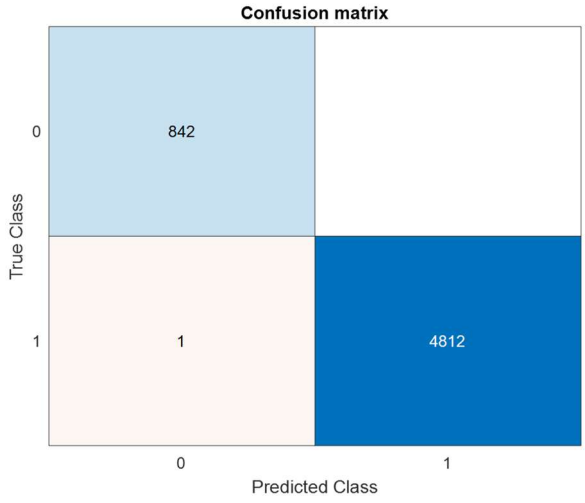

Table 5 Confusion Matrix Table (Test)

From the above table 6 of TRAIN-SET of IEEE-118 BUS, it is observed that there are 389 faults are correctly classified out of 390 faults and we can also detect that there are 2289 cyber-attacks

are correctly identified out of 2289 cyber-attacks.

From the above table 7 of TEST-SET of IEEE-118 BUS, it is observed that there are 842 faults are correctly classified out of 842 faults and we can also detect that there are 4812 cyber-attacks are correctly identified out of 4813 cyber-attacks.

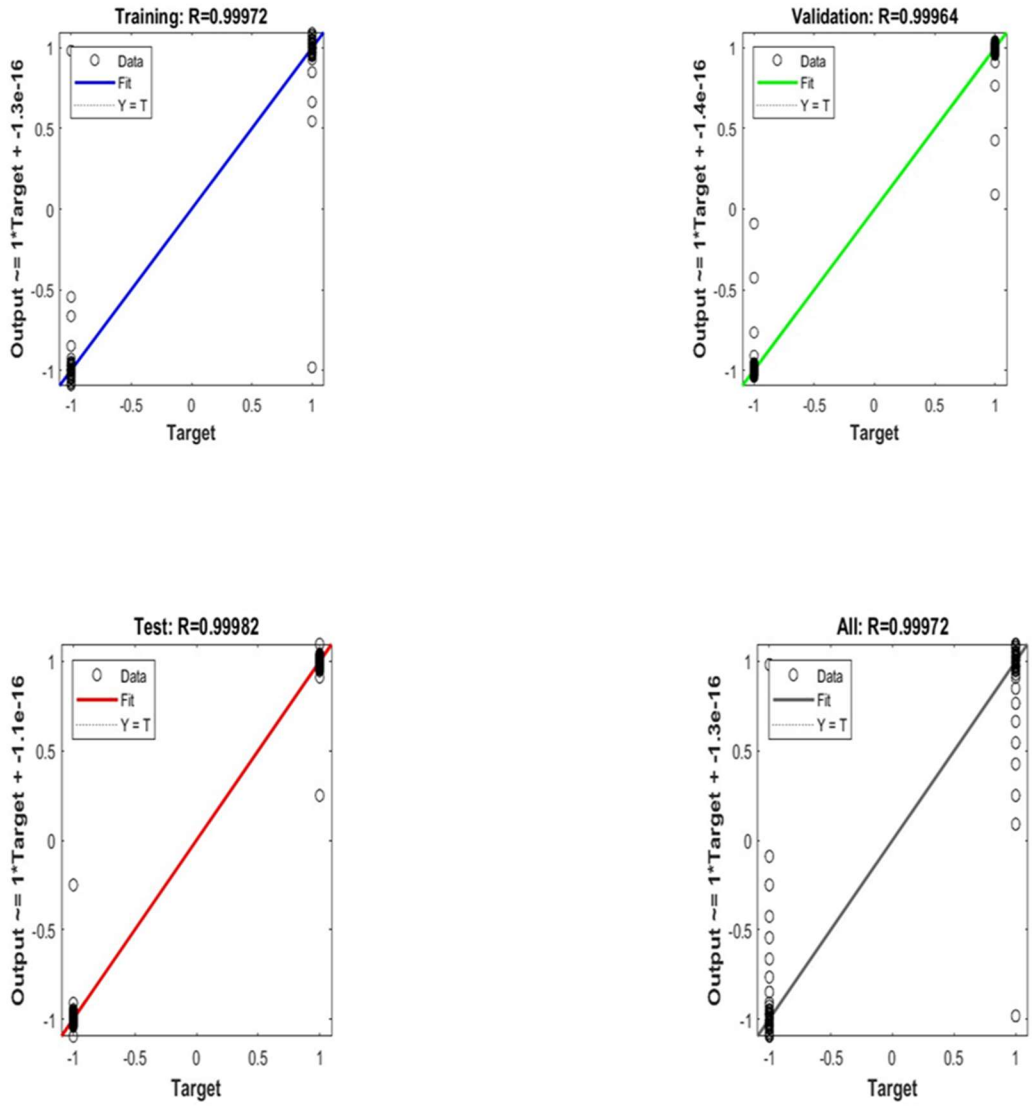

Fig.4: Regression value

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

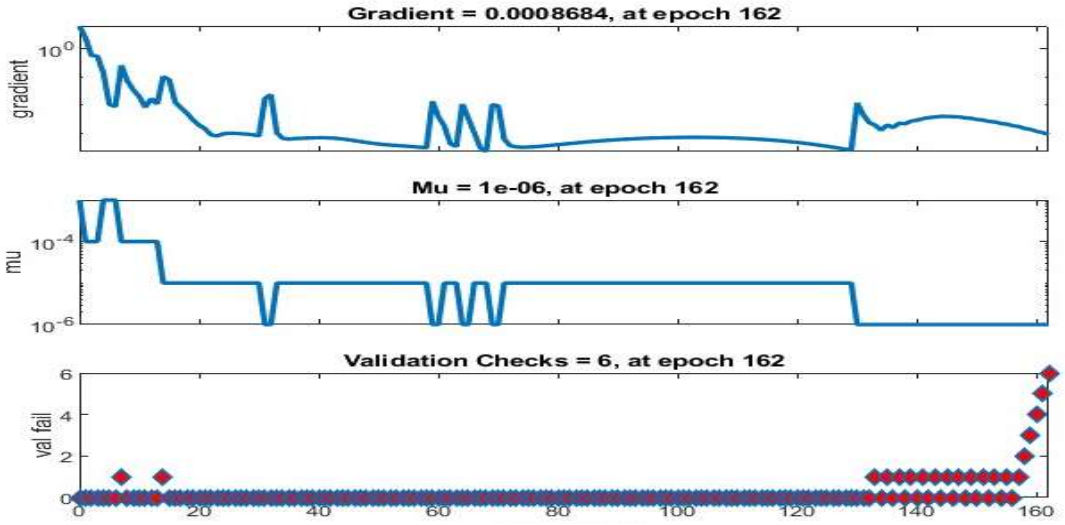

Fig. 5: Training stage

Above Fig.4 shows that Regression value is near to 1 so we can say that model is fit to particular target. In the given Fig. 5, validation check is 6 at epoch 162 which means model is struggling to generalize and make accurate prediction. High value of validation failure rate means model is overfitting.

4.2 Performance Metrics Comparison Of IEEE-14 Bus vs Benchmark

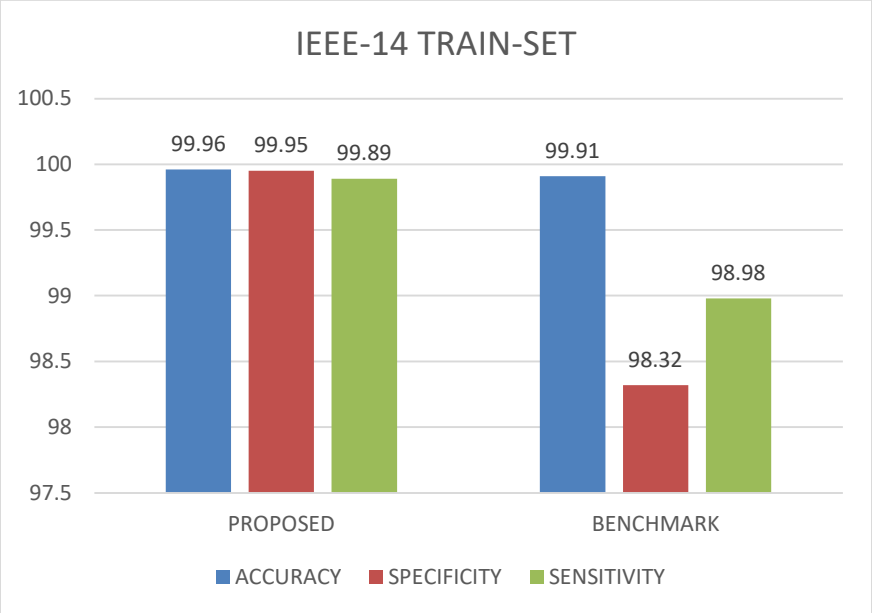

Fig. 6: IEEE-14 Bus Train-set

The Fig. 6 represents the TRAIN-SET analysis of IEEE-14 BUS with the x-axis being the proposed and Benchmark algorithm and the y-axis will be accuracy, Specificity, and Sensitivity

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

values. The estimation accuracy in this scenario is roughly 99.96% for faults and cyberattacks It shows that the Proposed algorithm produces effective results when compared with Benchmark algorithm.

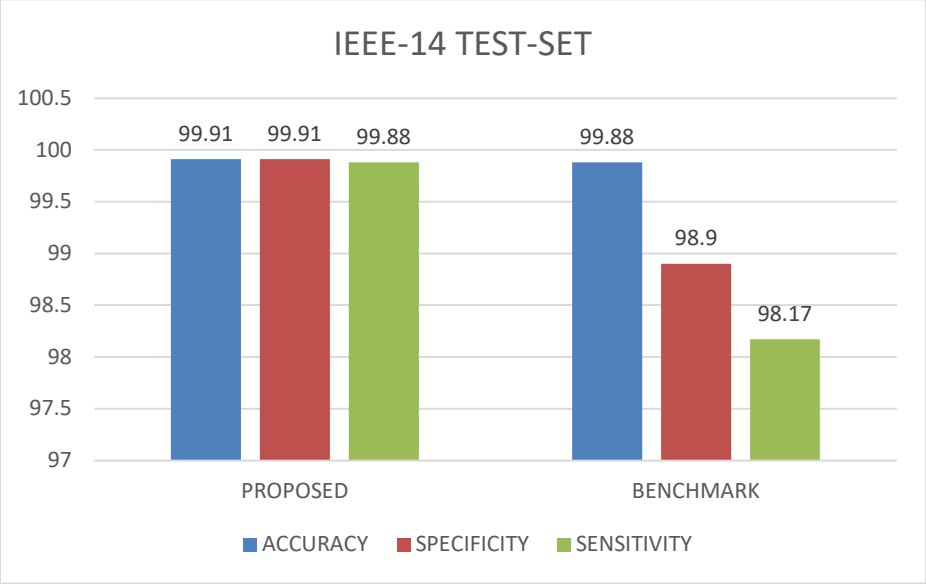

Fig. 7: IEEE-14 Bus Test-Set

The Fig. 7 represents the TEST-SET analysis of IEEE-14 BUS with the x-axis being the proposed and Benchmark algorithm and the y-axis will be accuracy, Specificity, and Sensitivity values. The estimation accuracy in this scenario is roughly 99.91% for faults and cyberattacks. It shows that the Proposed algorithm produces effective results when compared with Benchmark algorithm.

Performance Metrics Comparison Of IEEE-118 Bus vs Benchmark

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

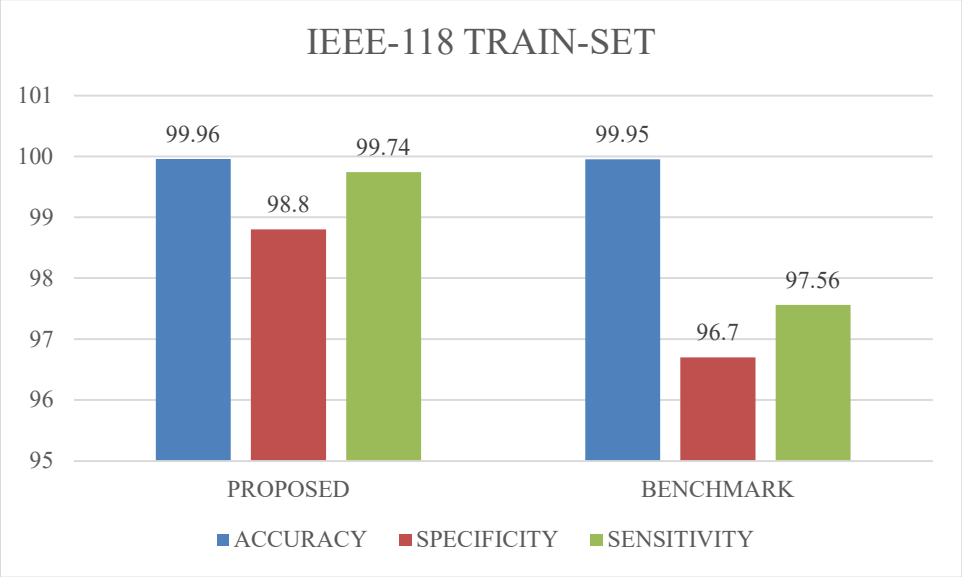

Fig. 8: IEEE-118 Bus Train-Set

The Fig. 8 represents the TRAIN-SET analysis of IEEE-118 BUS with the x-axis being the proposed and Benchmark algorithm and the y-axis will be accuracy, Specificity, and Sensitivity values. The estimation accuracy in this scenario is roughly 99.96% for faults and cyberattacks It shows that Proposed algorithm produces effective results when compared with Benchmark algorithm.

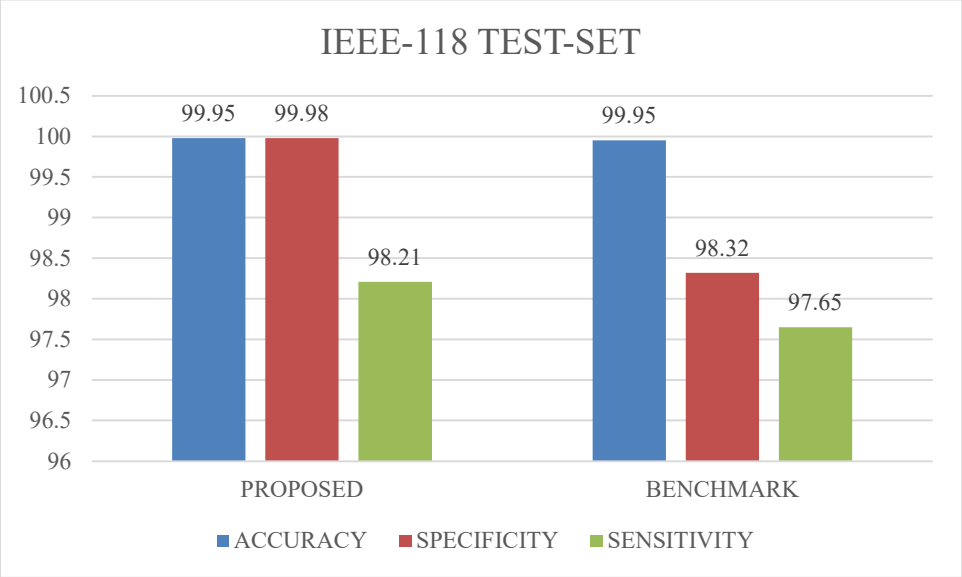

Fig. 9: IEEE-118 Bus Test-Set

The above Fig. 9 represents the TEST-SET analysis of IEEE-118 BUS with the x-axis being the proposed and Benchmark algorithm and the y-axis will be accuracy, Specificity, and Sensitivity

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

values. The estimation accuracy in this scenario is roughly 99.95% for faults and cyberattacks. It shows that proposed algorithm produces effective results when compared with Benchmark algorithm.

5. CONCLUSION

A Deep-reinforcement learning system is proposed in this article as a method for identifying cyber-attacks in distant relays. Due to the scattered nature of distance relays, which are extensively employed in the power grid to defend against failures, these relays are susceptible to cyberattacks despite their widespread usage. The algorithm that has been suggested makes use of a number of different agents in order to discover the most effective strategy for detecting cyberattacks in distant relays. Techniques like as deep reinforcement learning and supervised learning are used throughout each agent's education. The agents learn how to recognize assaults by monitoring the present state of the system and carrying out activities that aim to maximize the value of a reward function. The created technique has been used in case study, which include electricity grids with IEEE 14-bus bus configuration respectively. The incentive function was developed to optimize the detection accuracy of the agents while simultaneously reducing the number of false positives. The suggested approach is evaluated based on its performance on a standard dataset consisting of simulated cyberattacks. The findings demonstrate that the suggested algorithm performs better than the current methods in terms of the accuracy of attack detection as well as the rate of false alarms. It is anticipated that the method that has been suggested would be effective in improving the safety of the electric power grid's distance relays.

REFERENCES

1. Xing, K., Li, A., Jiang, R., & Jia, Y. (2021). Detection and Defense Methods of Cyber Attacks. MDATA: A New Knowledge Representation Model: Theory, Methods and Applications, 185-198.
2. Sarker, I. H. (2021). Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. SN Computer Science, 2(6), 420.Atallah, R. F., Assi, C. M., & Khabbaz, M. J. (2019). Scheduling the Operation of a Connected Vehicular Network Using Deep Reinforcement Learning. *IEEE Transactions on Intelligent Transportation Systems*, *20*(5), 1669–1682. https://doi.org/10.1109/TITS.2018.2832219
3. Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., ... & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. Journal of big Data, 8, 1-74.Batina, L., Picek, S., & Mondal, M. (Eds.). (2020). *Security, Privacy, and Applied Cryptography Engineering: 10th International Conference, SPACE 2020, Kolkata, India, December 17–21, 2020, Proceedings* (Vol. 12586). Springer International Publishing. https://doi.org/10.1007/978-3-030- 66626-2
4. Mayfield, K. P., Petty, M. D., & Whitaker, T. S. (2019). Machine Learning Cyberattack

Strategies with Petri Nets with Players, Strategies, and Costs. National Cyber Summit (NCS) Research Track, 1055, 232.

5. Burgos-Mellado, C., Zuñiga-Bauerle, C., Muñoz-Carpintero, D., Arias-Esquivel, Y., Càrdenas- Dobson, R., DragiČević, T., Donoso, F., & Watson, A. (2023). Reinforcement Learning-Based

   Method to Exploit Vulnerabilities of False Data Injection Attack Detectors in Modular Multilevel Converters. *IEEE Transactions on Power Electronics*, 1–15. https://doi.org/10.1109/TPEL.2023.3263728
6. CengiZ, E., & Gök, M. (2023). Reinforcement Learning Applications in Cyber Security: A Review.

   Sakarya University Journal of Science. https://doi.org/10.16984/saufenbilder.1237742
7. Chen, C., Cui, M., Fang, X., Ren, B., & Chen, Y. (2020). Load-altering attack-tolerant defense strategy for load frequency control system. *Applied Energy*, *280*, 116015. https://doi.org/10.1016/j.apenergy.2020.116015
8. Chen, W., Qiu, X., Cai, T., Dai, H.-N., Zheng, Z., & Zhang, Y. (2021). Deep Reinforcement Learning for Internet of Things: A Comprehensive Survey. *IEEE Communications Surveys & Tutorials*, *23*(3), 1659–1692. https://doi.org/10.1109/COMST.2021.3073036
9. Dehghan, M., Sadeghiyan, B., Khosravian, E., Moghaddam, A. S., & Nooshi, F. (2022). *ProAPT: Projection of APT Threats with Deep Reinforcement Learning* (arXiv:2209.07215). arXiv. http://arxiv.org/abs/2209.07215
10. Fang, D., Guan, X., Hu, B., Peng, Y., Chen, M., & Hwang, K. (2021). Deep Reinforcement Learning for Scenario-Based Robust Economic Dispatch Strategy in Internet of Energy. *IEEE Internet of Things Journal*, *8*(12), 9654–9663. https://doi.org/10.1109/JIOT.2020.3040294
11. Garrad, P., & Unnikrishnan, S. (2023). Reinforcement learning in VANET penetration testing.

    Results in Engineering, *17*, 100970. https://doi.org/10.1016/j.rineng.2023.100970
12. Hu, C., Yan, J., & Liu, X. (2022). Reinforcement Learning-Based Adaptive Feature Boosting for Smart Grid Intrusion Detection. *IEEE Transactions on Smart Grid*, 1–1. https://doi.org/10.1109/TSG.2022.3230730
13. Jin, Z., Yu, P., Guo, S. Y., Feng, L., Zhou, F., Tao, M., Li, W., Qiu, X.-S., & Shi, L. (2020). Cyber-

    Physical Risk Driven Routing Planning with Deep Reinforcement-Learning in Smart Grid Communication Networks. *2020 International Wireless Communications and Mobile Computing (IWCMC)*, 1278–1283. https://doi.org/10.1109/IWCMC48107.2020.9148342

14. Maeda, R., & Mimura, M. (2021). Automating post-exploitation with deep reinforcement learning.

Computers & Security, *100*, 102108. https://doi.org/10.1016/j.cose.2020.102108

15. Mern, J., Hatch, K., Silva, R., Brush, J., & Kochenderfer, M. J. (2021). *Reinforcement Learning for Industrial Control Network Cyber Security Orchestration* (arXiv:2106.05332). arXiv. http://arxiv.org/abs/2106.05332
16. Modirrousta, M. H., Arani, P. F., & Shoorehdeli, M. A. (2022). *Analysis of Anomalous Behavior in Network Systems Using Deep Reinforcement Learning with CNN Architecture* (arXiv:2211.16304). arXiv. http://arxiv.org/abs/2211.16304
17. Muhati, E., & Rawat, D. B. (2021). Asynchronous Advantage Actor-Critic (A3C) Learning for Cognitive Network Security. *2021 Third IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA)*, 106–113.

https://doi.org/10.1109/TPSISA52974.2021.00012

18. Nguyen, T. T., & Reddi, V. J. (2021). Deep Reinforcement Learning for Cyber Security. *IEEE Transactions on Neural Networks and Learning Systems*, 1–17. https://doi.org/10.1109/TNNLS.2021.3121870
19. Peng, Y., Liu, C., Liu, S., Liu, Y., & Wu, Y. (2022). SmartTRO: Optimising topology robustness for the Internet of Things via deep reinforcement learning with graph convolutional networks. *Computer Networks*, *218*, 109385. https://doi.org/10.1016/j.comnet.2022.109385
20. Wang, Y., Qiu, D., & Strbac, G. (2022). Multi-agent deep reinforcement learning for resilience- driven routing and scheduling of mobile energy storage systems. *Applied Energy*, *310*, 118575. https://doi.org/10.1016/j.apenergy.2022.118575
21. Rajaee, M., & Mazlumi, K. (2023). Multi-Agent Distributed Deep Learning Algorithm to Detect Cyber-Attacks in Distance Relays. *IEEE Access*, *11*, 10842–10849. https://doi.org/10.1109/ACCESS.2023.3239684
22. Zhang, T., Xu, C., Zhang, B., Shen, J., Kuang, X., & Grieco, L. A. (2022). Toward Attack-Resistant Route Mutation for VANETs: An Online and Adaptive Multi-agent Reinforcement Learning Approach. *IEEE Transactions on Intelligent Transportation Systems*, *23*(12), 23254–23267. https://doi.org/10.1109/TITS.2022.3198507
23. Yang, L., Tao, J., Liu, Y.-H., Xu, Y., & Su, C.-Y. (2023). Energy scheduling for DoS attack over multi-hop networks: Deep reinforcement learning approach. *Neural Networks*, *161*, 735–745. https://doi.org/10.1016/j.neunet.2023.02.028
24. Yu, L., Gao, Z., Qin, S., Zhang, M., Shen, C., Guan, X., & Yue, D. (2020). *Deep Reinforcement Learning for Smart Grid Protection Against Coordinated Multistage Transmission Line Attacks* (arXiv:2011.14526). arXiv. http://arxiv.org/abs/2011.14526
25. Rezwan, S., & Choi, W. (2021). A Survey on Applications of Reinforcement Learning in

Flying Ad-Hoc Networks. *Electronics*, *10*(4), 449. https://doi.org/10.3390/electronics10040449

26. Stanly Jayaprakash, J., Priyadarsini, M. J. P., Parameshachari, B. D., Karimi, H. R., & Gurumoorthy, S. (2022). Deep *Q* -Network with Reinforcement Learning for Fault Detection in Cyber-Physical Systems. *Journal of Circuits, Systems and Computers*, *31*(09), 2250158. https://doi.org/10.1142/S0218126622501584
27. Alavizadeh, H., Alavizadeh, H., & Jang-Jaccard, J. (2022). Deep Q-learning based reinforcement learning approach for network intrusion detection. Computers, 11(3), 41.Shafiq, M., Tian, Z., Sun, Y., Du, X., & Guizani, M. (2020). Selection of effective machine learning algorithm and Bot-IoT attacks traffic identification for the internet of things in a smart city. *Future Generation Computer Systems*, *107*, 433–442. https://doi.org/10.1016/j.future.2020.02.017
28. Sun, Z., Wang, N., Lin, H., & Zhou, X. (2023). Persistent coverage of UAVs based on deep reinforcement learning with wonderful life utility. *Neurocomputing*, *521*, 137–145. https://doi.org/10.1016/j.neucom.2022.11.091
29. Wan, X., Zeng, L., & Sun, M. (2022). Exploring the Vulnerability of Deep Reinforcement Learning-based Emergency Control for Low Carbon Power Systems. *Proceedings of the Thirty- First International Joint Conference on Artificial Intelligence*, 3954–3961. https://doi.org/10.24963/ijcai.2022/549
30. Elharrouss, O., Akbari, Y., Almaadeed, N., & Al-Maadeed, S. (2022). Backbones-review: Feature extraction networks for deep learning and deep reinforcement learning approaches. arXiv preprint arXiv:2206.08016.
31. Wu, D., Kalathil, D., Begovic, M. M., Ding, K. Q., & Xie, L. (2022). Deep Reinforcement Learning-Based Robust Protection in DER-Rich Distribution Grids. *IEEE Open Access Journal of Power and Energy*, *9*, 537–548. https://doi.org/10.1109/OAJPE.2022.3161904